PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Severe varicose veins and the risk of mortality: A nationwide population-based cohort study
AUTHORS	Wu, Nan-Chun; Chen, Zhih-Cherng; Feng, I-Jung; Ho, Chung-Han; Chiang, Chun-Yen; Wang, Jhi-Joung; Chang, Wei-Ting

VERSION 1 – REVIEW

REVIEWER	Arina J. ten Cate-Hoek Maastricht University Medical Centre, the Netherlands.
REVIEW RETURNED	08-Oct-2019

GENERAL COMMENTS	Overall well written and interesting paper. However there are some major concerns. Although scenario's to elucidate the associations found are described, altogether the paper induces many questions. The main concerns to me are that it is not clear whether propensity score matching was successful, and whether the crude diagnostic method and sub-categorisation for vv is entirely appropriate. Overall the hypothesis is of interest, vv is a common condition and indirect associations with cardiovascular disease and mortality have been shown earlier. (vv as risk factor for DVT, DVT as risk factor for mortality) Setting: population based cohort study with retrospectively acquired data. Strengths: large sample size, ascertainment of diagnosis by specialist Weaknesses: retrospective study based on claims data, missing information important to the outcome such as smoking history, mobility, history of pregnancy, BMI and DM disease control (glycated haemoglobin levels), inadequate testing for imbalance of the model.1 It was stated in the bullet points that age and sex were not known, however in Table 1 + 2. Age and sex are presented. This might be an error. 1Propensity scores are used to balance two non-equivalent groups on observed characteristics to obtain less biased estimates. Two important shortcomings concerning the use of propensity score matching in this study can be identified. 1. The model should include all variables that affect the outcome (there should be no unmeasured confounders), however the model in this study does lack information on a number of important confounding variables as stated by the authors. 2."Proof" that balance between the groups studied is achieved by

	propensity score matching is shown in the baseline table. Table 1 as well as Table 2, however show hypothesis testing, which has been criticized as optimal means of establishing that there is balance in the analytic sample. The method suggested to test for this balance is to compute the standardized difference of each covariate. The general agreement being that differences <0.1 are negligible. Whenever this is not the case for a number of covariates the conclusion should be that the model has not been specified correctly. By comparing the proportions between groups using statistical significance testing one does not obtain information on the balance of covariates in the actual sample. The grading system used was a crude grading: grade 1. vv without inflammation or ulceration, grade 2 vv with either inflammation or ulceration, grade 3 vv with ulceration and inflammation. This does not correspond with the CEAP classification. A vascular specialist is described to have examined the records, and to have compared CEAP stages with vv grading. It is however not clear from the paper what the outcome of this comparison was. KM curves were used to assess 3,6 and 9-year survival rates. There was a significant difference in survival between vv patients (grade II and III) and those without. Risk of mortality overall was 1.26 times higher, however there was effect modification by age and gender showing larger effect in those <65 and in male patients. It is not discussed extensively what the meaning of this finding could be. Ascertainment of death was well established, however ascertainment of the other outcomes is not clearly described.

REVIEWER	Jürgen Prochaska University Medical Center Mainz, Germany
REVIEW RETURNED	08-Oct-2019

GENERAL COMMENTS	In the manuscript entitled "Varicose Veins are Associated with Mortality and Cardiovascular Events: A Nationwide Cohort Study" Wu and coauthors investigate the clinical relevance of varicose veins. - One of the major limitations of propensity score matching is the quality of underlying data assessment, i.e. matching is only as good as the assessment of factors used. Besides this potential source of bias I would be interested in the selection of variables used for propensity score matching since only selected risk factors and comorbidities were selected. - The grading of venous veins was done by a reviewer. I would be interested about the comparability of this grading with CEAP classification scheme and in particular about the distribution of CEAP classes in the study sample. - For the analytical outcome analyses I wondered about the selection of variables (e.g. cancer status was not considered).

REVIEWER	Yan Ren Lin Changhua Christian Hospital, Changhua,Taiwan.
REVIEW RETURNED	26-Oct-2019

GENERAL COMMENTS	This paper is generally well written. The analysis is also appropriate. However, the objective is not specific. “Cardiovascular events” is too much and limited the clinical applications. It is difficult to follow.

	I suggest that authors should focus on a specific area/disease (for example, ACS or aorta dissection or other targeted diseases, just choose one). Finally, make strong association between VV and your targeted diseases.

REVIEWER	Ana Timoteo Santa Marta Hospital, Lisbon, Portugal
REVIEW RETURNED	11-Nov-2019

GENERAL COMMENTS	I have been asked to review only the statistical part of this paper. This study was performed with data from the Taiwan´s National Health Insurance Database. The authors identified 4,807 patients with newly diagnosed varicose veins between 1999 and 2012 and they selected a matched-control group of 35,456 individuals. Their main objective was to investigate the factors associated with overall mortality in patients with varicose veins, specifically if varicose veins were associated with the occurrence of all-cause mortality and major cardiovascular events. From a statistical point of view, the selection of matched controls for assessment in a propensity score matching technique is a very useful technique to use as an alternative to randomization. It can use retrospective data in a similar way as in randomized controlled clinical trials, allowing the study of more individuals. For that reason, sample size in this study is quite good. The matching technique used by the authors is also appropriate. The variable description and comparison is also adequate. The multivariate regression analysis used by the authors is Conditional Cox Proportional hazards regression analysis, also adequate, together with Kaplan-Meier curves and Log-rank test for comparison. Sub-groups analysis was also performed. There are however some important issues with the methods used by the authors: 1 – They performed a validation of the accuracy of varicose veins diagnosis and CEAP grading. However, no validation was done for the outcomes. The adjudication of events was not validated and this is an important source of bias. Also, over the years, definitions changed not only for outcomes but also for the variables included in the adjustment for multivariate analysis. I think there were no strict definitions for hypertension, diabetes, hyperlipidemia. Also the variable CAD, what is the definition? In my opinion, a patient with a previous history of myocardial infarction, stroke, coronary angioplasty or CABG should have been excluded from the study and it does not seem to be the case. 2 – Not all possible predictors of outcome were included in the multivariate analysis. They only included data available from the database but there are probably other variables that the authors could not identify and that is another cause for possible bias in the study.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1
Reviewer Name: Arina J. ten Cate-Hoek
Institution and Country: Maastricht University Medical Centre, the Netherlands.
Please state any competing interests or state ‘None declared’: None declared

Question 1. Overall well written and interesting paper. However there are some major concerns. Although scenario's to elucidate the associations found are described, altogether the paper induces many questions. The main concerns to me are that it is not clear whether propensity score matching was successful, and whether the crude diagnostic method and sub-categorisation for vv is entirely appropriate.

Reply 1. We appreciate reviewer's thoughtful comments. The study design has been modified and the standardized mean difference (SMD) of each covariate has been calculated to exam the balance between VV and matched control groups.

Three groups of patients with varicose vein (VV) were separately matched with individuals in control group at a 1:4 ratio based on age, sex, hypertension, diabetes, hyperlipidemia, coronary artery disease and history of pregnancy. The baseline characteristics and comorbid medical disorders for 3 grading VV groups and 3 separately matched controls were updated in supplement Table 2 and the balance of matched characteristics between control and VV groups were evaluated by SMD. Three VV groups and matched controls were separately combined into one overall VV group and one overall control group. The difference of the baseline characteristics were also checked by SMD and displayed in Table 1. A SMD smaller than 0.1 indicates an adequate balance between VV and control groups.

The result was addressed on Page 10 as the follows,

"Lower survival rates over time were observed in patients with highest VV severity (grades 3) but not in those with grade 1-2. Significant difference between survival curves between VV grading 3 and corresponding controls were revealed by log rank test ($p < 0.0001$). However, there is no significant difference were found between survival curves of patients with VV severity grades (1-2) and corresponding controls (grade 1: $p = 0.3191$; grade 2: $p=0.3599$)."

Question 2. Overall the hypothesis is of interest, vv is a common condition and indirect associations with cardiovascular disease and mortality have been shown earlier. (vv as risk factor for DVT, DVT as risk factor for mortality)

Setting: population based cohort study with retrospectively acquired data.

Strengths: large sample size, ascertainment of diagnosis by specialist

Weaknesses: retrospective study based on claims data, missing information important to the outcome such as smoking history, mobility, history of pregnancy, BMI and DM disease control (glycated haemoglobin levels), inadequate testing for imbalance of the model.1

Reply 2. We appreciate your comments while lacking the above mentioned potential confounding factors is a limitation of this cohort.

For a more solid result, we controlled additionally 6 confounding factors, such as chronic obstructive pulmonary disease (ICD9 code:490-496) , cancer (ICD9 code:140-208), atrial fibrillation (ICD9 code:427.31), heart failure (ICD9 code:428), ischemic heart disease (ICD9 code:410-414), chronic renal insufficiency (ICD9 code:403, 404, 582, 585-588), and the results in table 2, 3 and 4 is updated.

Further, as suggested, association evaluation between VV and mortality and MACE were separately additionally calculated by including history of pregnancy (ICD-9 cods: V22, V23.2, 761.5) before index date and results showed that VV significantly increased 1.369 times risk of mortality and 2.055 times risk of MACE. For the Grade 3 VV population, 1.838 and 2.456 times risk of mortality and MACE separately found, by comparing with matched control (details were displayed in Supplement Table 5).

However, NHIRD is lacking lifestyle factors and laboratory data, such as smoking history, mobility, BMI and DM disease control (glycated haemoglobin levels).

To overcome the limitation, we also performed analysis of unmeasured confounders. Regarding the potential effect, smoking, on the estimation of association between VV and mortality, we performed external adjustment for analysis of unmeasured confounders [1]. According to the latest meta-analysis [2], smoking people have 1.70 greater death risk than nonsmokers in Asia. The estimated smoking prevalence among control group is 21.1%. For VV population, 30% smoking prevalence were assumed. For the included population, the crude OR is 1.44 and the smoking adjusted OR is 1.36 calculated by Greenland's method [1]. For VV at grading 3 and matched control, the crude OR is 2.03 and the smoking adjusted OR is 1.93.

1. Rothman, K.J., S. Greenland, and T. L. Lash., Modern Epidemiology. 3rd ed ed. 2008, Philadelphia: Lippincott Williams & Wilkins.
2. Yang JJ, Yu D, Wen W, et al. Tobacco smoking and mortality in Asia: a pooled meta-analysis. JAMA Netw Open 2019; 2: e191474.

Question 3. It was stated in the bullet points that age and sex were not known, however in Table 1 + 2. Age and sex are presented. This might be an error.
Reply 3. We apologize for the typo. The description of “age and sex” were removed The bullet point was listed as the follows, However, some risk factors of varicose vein including smoking habits, lack of movement, overweight and glycated hemoglobin levels were not available in this database.”

Question 4.Propensity scores are used to balance two non-equivalent groups on observed characteristics to obtain less biased estimates. Two important shortcomings concerning the use of propensity score matching in this study can be identified. 1. The model should include all variables that affect the outcome (there should be no unmeasured confounders), however the model in this study does lack information on a number of important confounding variables as stated by the authors.
Reply 4. Thanks for reviewer’s comments. The associations between population with VV and without VV for the all-cause mortality and MACE in Table 2, 3 and the association within each VV severity groups (in Table 4) were updated under adjustment of 6 additional confounding factors; chronic obstructive pulmonary disease (ICD9 code:490-496) , cancer (ICD9 code:140-208), atrial fibrillation (ICD9 code:427.31), heart failure (ICD9 code:428), ischemic heart disease (ICD9 code:410-414), chronic renal insufficiency (ICD9 code:403, 404, 582, 585-588). However, as described in Reply 2 , association evaluation between VV and mortality and MACE were separately additionally calculated by including history of pregnancy (ICD-9 cods: V22, V23.2, 761.5) before index date and results showed that VV significantly increased 1.369 times risk of mortality and 2.055 times risk of MACE (details were displayed in Supplement Table 5). However, NHIRD is lacking lifestyle factors and laboratory data, such as smoking history, mobility, BMI and DM disease control (glycated haemoglobin levels). We will be extremely grateful if you may understand the limitation of this study while it remains showing important information regarding the possible under-evaluation of the risks of VVs.

Question 5. “Proof” that balance between the groups studied is achieved by propensity score matching is shown in the baseline table. Table 1 as well as Table 2, however show hypothesis testing, which has been criticized as optimal means of establishing that there is balance in the analytic sample. The method suggested to test for this balance is to compute the standardized difference of each covariate. The general agreement being that differences <0.1 are negligible. Whenever this is not the case for a number of covariates the conclusion should be that the model has not been specified correctly. By comparing the proportions between groups using statistical significance testing one does not obtain information on the balance of covariates in the actual sample.
Reply 5. We appreciate reviewer’s comments. Similar to Reply 1, we modified the study design and the standardized mean difference (SMD) of each covariate has been calculated to exam the balance between VV and matched control groups. Briefly, three groups of patients with varicose vein (VV) were separately matched with individuals in control group at a 1:4 ratio based on age, sex, hypertension, diabetes, hyperlipidemia, coronary artery disease and history of pregnancy. The baseline characteristics and comorbid medical disorders for 3 grading VV groups and 3 separately matched controls were updated in supplement Table 2 and the balance of matched characteristics between control and VV groups were evaluated by SMD. Three VV groups and matched controls were separately combined into one overall VV group and one overall control group. The difference of the baseline characteristics were also checked by SMD and displayed in Table 1. A SMD smaller than 0.1 indicates an adequate balance between VV and control groups. The result was addressed on Page 10.

Question 6.The grading system used was a crude grading: grade 1. vv without inflammation or ulceration, grade 2 vv with either inflammation or ulceration, grade 3 vv with ulceration and inflammation. This does not correspond with the CEAP classification. A vascular specialist is described to have examined the records, and to have compared CEAP stages with vv grading. It is however not clear from the paper what the outcome of this comparison was.
Reply 6. Thanks for your comment. To evaluate the accuracy of VV diagnosis and the association between CEAP classification and our ICD-9-CM–Derived VV Grading, a vascular specialist reviewed patients' clinical records and compared CEAP stages with our ICD-9-CM–derived grades in inpatients. Notably, compared with CEAP stage, as determined based on chart reviews, only a few inpatients were incorrectly or unclearly diagnosed using ICD-9-CM–derived VV codes (Supplement Table 3). The sensitivity and specificity of ICD-9-CM–derived grading were up to 95.2% and 97.6%, respectively. Thus, we added the description "In addition to examining the accuracy of VV diagnosis, the reviewer compared CEAP stages with our ICD-9-CM–derived grades in inpatients." In the section of Method.
Furthermore, a 0.918 (95%CI = [0.878, 0.957]) kappa score between CEAP stages and grading severity is calculate. This value is used to show the consistency between these two evaluation methods. Kappa score between 0.81 and 1.00 presents an almost perfect agreement. Therefore, it is appropriate to present CEAP stages by severity grading.

Question 7. KM curves were used to assess 3, 6 and 9-year survival rates. There was a significant difference in survival between vv patients (grade II and III) and those without. Risk of mortality overall was 1.26 times higher, however there was effect modification by age and gender showing larger effect in those <65 and in male patients. It is not discussed extensively what the meaning of this finding could be.
Reply 7. Thanks for your comment. It is notably that usually older age and female sex are associated with an increasing prevalence of VV. However, Heit et al. have reported that younger patients with VV were at a significantly increased risk of subsequent DVT (new Ref 20). Also, we added the Ref discussing a phenomenon that although a lower grade of VV (CEAP 2-3) has been observed in 50.5% of females and in 30.1% of males, a higher grade of VV with trophic skin changes (CEAP 4-6) were found in 2.8% of females and 5.4% of males. Also, DVT was more common in males compared with females (11.3% vs 7.8%). We hypothesized that the earlier onset of VV in the younger population implies a higher risk of concomitant arterial diseases or systemic inflammations. As described previously, female sex, pregnancy, and predominately being in the sitting posture are risk factors for VV.20 However, despite the valid correlation between use of estrogen supplements and DVT, whether sex hormones contribute to the development of VV remains unclear. The revised discussion was addressed on page 13 as the follows,
"Similarly, Lohr et al also reviewed that Although a lower grade of VV (CEAP 2-3) has been observed in 50.5% of females and in 30.1% of males, a higher grade of VV with trophic skin changes (CEAP 4-6) were found in 2.8% of females and 5.4% of males. Also, DVT was more common in males compared with females (11.3% vs 7.8%).20 Earlier onset of VV in the younger population implies a higher risk of concomitant arterial diseases or systemic inflammations. As described previously, female sex, pregnancy, and predominately being in the sitting posture are risk factors for VV.21 However, despite the valid correlation between use of estrogen supplements and DVT, whether sex hormones contribute to the development of VV remains unclear." We also added Ref 20.

Question 8. Ascertainment of death was well established, however ascertainment of the other outcomes is not clearly described.
Reply 8. Thanks for your comments. A recently published article entitled. "Taiwan ' s National Health Insurance Research Database: past and future" evaluated and supported the accuracy of several major outcomes, including myocardial infraction, hypertension, diabetes, stroke, heart failure and varicose vein, in Taiwan National Health Insurance Research Database (NHIRD). We cited this article

as Ref 22.

Reviewer: 2
Reviewer Name: Jürgen Prochaska
Institution and Country: University Medical Center Mainz, Germany
Please state any competing interests or state 'None declared': None declared.

Question 1. One of the major limitations of propensity score matching is the quality of underlying data assessment, i.e. matching is only as good as the assessment of factors used. Besides this potential source of bias I would be interested in the selection of variables used for propensity score matching since only selected risk factors and comorbidities were selected.
Reply 1. We appreciate reviewer's comments. The study design have been modified and the standardized mean difference (SMD) of each covariate has been calculated to exam the balance between VV and matched control groups.
Three groups of patients with varicose vein (VV) were separately matched with individuals in control group at a 1:4 ratio based on age, sex, hypertension, diabetes, hyperlipidemia, coronary artery disease and history of pregnancy. The baseline characteristics and comorbid medical disorders for 3 grading VV groups and 3 separately matched controls were updated in supplement Table 2 and the balance of matched characteristics between control and VV groups were evaluated by SMD. Three VV groups and matched controls were separately combined into one overall VV group and one overall control group. The difference of the baseline characteristics were also checked by SMD and displayed in Table 1. A SMD smaller than 0.1 indicates an adequate balance between VV and control groups.
The result was addressed on Page 10 as the follows,
"Lower survival rates over time were observed in patients with highest VV severity (grades 3) but not in those with grade 1-2. Significant difference between survival curves between VV grading 3 and corresponding controls were revealed by log rank test ($p < 0.0001$). However, there is no significant difference were found between survival curves of patients with VV severity grades (1-2) and corresponding controls (grade 1: $p = 0.3191$; grade 2: $p=0.3599$)."

Question 2.The grading of venous veins was done by a reviewer. I would be interested about the comparability of this grading with CEAP classification scheme and in particular about the distribution of CEAP classes in the study sample.

Reply 2. Thanks for your comment. To evaluate the accuracy of VV diagnosis and the association between CEAP classification and our ICD-9-CM–Derived VV Grading, a vascular specialist reviewed patients' clinical records and compared CEAP stages with our ICD-9-CM–derived grades in inpatients. Notably, compared with CEAP stage, as determined based on chart reviews, only a few inpatients were incorrectly or unclearly diagnosed using ICD-9-CM–derived VV codes (Supplement Table 3). The sensitivity and specificity of ICD-9-CM–derived grading were up to 95.2% and 97.6%, respectively. Thus, we added the description "In addition to examining the accuracy of VV diagnosis, the reviewer compared CEAP stages with our ICD-9-CM–derived grades in inpatients." In the section of Method.
Furthermore, a 0.918 (95%CI = [0.878, 0.957]) kappa score between CEAP stages and grading severity is calculate. This value is used to show the consistency between these two evaluation methods. Kappa score between 0.81 and 1.00 presents an almost perfect agreement. Therefore, it is appropriate to present CEAP stages by severity grading.

Question 3.For the analytical outcome analyses I wondered about the selection of variables (e.g. cancer status was not considered).
cancer (ICD-9 codes 140 – 208, concomitantly registered with catastrophic illness)
Reply 3. We appreciate reviewer's suggestion. Kappa score was calculated for evaluate the consistency between varicose veins and CEAP grading. Result show a 0.918 (95%CI = [0.878, 0.957]) value of kappa score, represented for an almost perfect agreement. As recommend by the reviewer, population with a history of myocardial infarction, stroke, coronary angioplasty or CABG were excluded and analyzed. The results shown that VV separately increased 1.36 (95% CI = [1.181, 1.567]) and 1.952 (95% CI = [1.799, 2.119]) times risk of mortality and MACE under the control of age, gender, hypertension, diabetes, hyperlipidemia, coronary artery disease, chronic obstructive pulmonary disease, cancer, atrial fibrillation, heart failure, ischemic heart disease and chronic renal insufficiency. Further, grade 3 VV displayed 1.842 and 2.316 greater risk of mortality and MACE than matched controls.

Reviewer: 3
Reviewer Name: Yan Ren Lin
Institution and Country: Changhua Christian Hospital, Changhua, Taiwan.
Please state any competing interests or state 'None declared': none

Question 1. Please leave your comments for the authors below
This paper is generally well written. The analysis is also appropriate. However, the objective is not specific. "Cardiovascular events" is too much and limited the clinical applications. It is difficult to follow. I suggest that authors should focus on a specific area/disease (for example, ACS or aorta dissection or other targeted diseases, just choose one). Finally, make strong association between VV and your targeted diseases.

Reply 1. We appreciate the reviewer's comment. Although the associated between VV and arterial or venous thrombosis has been reported, whether the severe VV correlates to mortality remains unknown. Thus, the aim of this study is to investigate the factors associated with overall mortality in patients with VV. We revised the title as the follows,
"The severe varicose veins and the risk of mortality: A nationwide population-based cohort study"

Reviewer: 4
Reviewer Name: Ana Timoteo
Institution and Country: Santa Marta Hospital, Lisbon, Portugal
Please state any competing interests or state 'None declared': None declared
Question 1. They performed a validation of the accuracy of varicose veins diagnosis and CEAP grading. However, no validation was done for the outcomes. The adjudication of events was not validated and this is an important source of bias. Also, over the years, definitions changed not only for outcomes but also for the variables included in the adjustment for multivariate analysis. I think there were no strict definitions for hypertension, diabetes, hyperlipidemia. Also the variable CAD, what is the definition? In my opinion, a patient with a previous history of myocardial infarction, stroke, coronary angioplasty or CABG should have been excluded from the study and it does not seem to be the case. ICD
Reply 1. We thanks for the reviewer's thoughtful suggestion. Kappa score was calculated for evaluate the consistency between varicose veins and CEAP grading. Result show a 0.918 (95%CI = [0.878, 0.957]) value of kappa score, represented for an almost perfect agreement. As recommend by the reviewer, population with a history of myocardial infarction, stroke, coronary angioplasty or CABG were excluded and analyzed. The results shown that VV separately increased 1.36 (95% CI = [1.181, 1.567]) and 1.952 (95% CI = [1.799, 2.119]) times risk of mortality and MACE under the control of age, gender, hypertension, diabetes, hyperlipidemia, coronary artery disease, chronic obstructive pulmonary disease, cancer, atrial fibrillation, heart failure, ischemic heart disease and chronic renal insufficiency. Further, grade 3 VV displayed 1.842 and 2.316 greater risk of mortality and MACE than matched controls. The associated description was addressed on Page 11.

Question 2. Not all possible predictors of outcome were included in the multivariate analysis. They only included data available from the database but there are probably other variables that the authors could not identify and that is another cause for possible bias in the study.
Reply 2. We appreciate your comments while lacking the above mentioned potential confounding factors is a limitation of this cohort. For a more solid result, we controlled additionally 6 confounding factors, such as chronic obstructive pulmonary disease (ICD9 code:490-496) , cancer (ICD9 code:140-208), atrial fibrillation (ICD9 code:427.31), heart failure (ICD9 code:428), ischemic heart disease (ICD9 code:410-414), chronic renal insufficiency (ICD9 code:403, 404, 582, 585-588), and the results in table 2, 3 and 4 is updated. Further, as suggested, association evaluation between VV and mortality and MACE were separately additionally calculated by including history of pregnancy (ICD-9 cods: V22, V23.2, 761.5) before index date and results showed that VV significantly increased 1.369 times risk of mortality and 2.055 times risk of MACE. For the Grade 3 VV population, 1.838 and 2.456 times risk of mortality and MACE separately found, by comparing with matched control (details were displayed in Supplement Table 5). However, NHIRD is lacking lifestyle factors and laboratory data, such as smoking history, mobility, BMI and DM disease control (glycated haemoglobin levels). To overcome the limitation, we also performed analysis of unmeasured confounders. Regarding the potential effect, smoking, on the estimation of association between VV and mortality, we performed external adjustment for analysis of unmeasured confounders [1]. According to the latest meta-analysis [2], smoking people have 1.70 greater death risk than nonsmokers in Asia. The estimated smoking prevalence among control group is 21.1%. For VV population, 30% smoking prevalence were assumed. For the included population, the crude OR is 1.44 and the smoking adjusted OR is 1.36 calculated by Greenland's method [1]. For VV at grading 3 and matched control, the crude OR is 2.03 and the smoking adjusted OR is 1.93.

1. Rothman, K.J., S. Greenland, and T. L. Lash., Modern Epidemiology. 3rd ed ed. 2008, Philadelphia: Lippincott Williams & Wilkins.
2. Yang JJ, Yu D, Wen W, et al. Tobacco smoking and mortality in Asia: a pooled meta-analysis. JAMA Netw Open 2019; 2: e191474.
We will be extremely grateful if you may understand the limitation of this study while it remains showing important information regarding the possible under-evaluation of the risks of VVs.

VERSION 2 – REVIEW

REVIEWER	ten Cate MUMC, the Netherlands
REVIEW RETURNED	01-Feb-2020

GENERAL COMMENTS	Last time I missed a rebuttal letter addressing the reviewers' remarks with the point to point answers. It was not clear at that time what had been changed. The rebuttal letter has made this clear. The authors have addressed all my queries to my satisfaction. However, before any publication the article needs thorough editing and language correction. At this point in time I do only have minor suggestions to improve the manuscript. --- Title: I would suggest to delete “The” as first word in the title. The strength and limitation section: Item 2. suggest to swap “should be” for “were” Pg. 10 Line 12: “However, there is no significant difference were found” Rephrase: However, no significant differences were found? Pg 10 lines 22-24 : Notably, despite no significant effect of VV on the survival of patients with hypertension, hyperlipidemia or coronary artery disease was observed?, patients with both VV and diabetes presented a 1.50 times higher risk of mortality compared with those without VV (adjusted HR: 1.50; 95% CI: 1.05–2.15; p = 0.0254). Pg 11. Line 3: under the control of = adjusted for? Pg 11, Line 12: In patients with VV, 3-, 6-, and 9-year MACE-free rates were 91.17% (91.2%), 84.99% (85%), and 79.27% (79.3%), Pg 11, lines 18-19: including DVT and PE (Grade 3: adjusted HR: 38.42 (38.4); 95% CI, 16.38 (16.4)–90.13(90.1); p < 0.0001) (Table 4). Pg 12, lines 10-11: The calculated kappa score between CEAP stages and grading severity is 0.918 (0.92) (95%CI = [0.878(0.88), 0.957(0.96)]). Pg 12, line 22: treated VV patients were found 1.36 times risks of mortality (adjusted HR (95%CI) = 1.362(136) (1.18, 1.57), p-value<0.0001) Page 14, lines 13-16: Please rephrase the following insert: “ Similarly, Lohr et al also reviewed that Although a lower grade of VV

	(CEAP 2-3) has been observed in 50.5% of females and in 30.1% of males, a higher grade of VV with trophic skin changes (CEAP 4-6) were found in 2.8% of females and 5.4% of males.20 Also, DVT was more common in males compared with females (11.3% vs 7.8%).20" The discussion could profit from editorial revision. Tables 2, 3,4. And supplement tables: Please remove third decimal throughout in the table. (1.431 (1.247, 1.643) should be 1.43 (1.25, 1.64) etc. Figure 1,2: There is a 12 year follow up , the KM plots show 15 years and 0 survivors. I would suggest limiting the KM plots to the observed follow up period of 12 years.

REVIEWER	Jürgen Prochaska University Medical Center Mainz Germany
REVIEW RETURNED	06-Jan-2020

GENERAL COMMENTS	In this version, I could not find a point by point response to the comments raised by reviewers although the manuscript has been overworked (at least in my online account at the journal). Please provide or resend for final review of this revision 1.

REVIEWER	Yan Ren Lin Changhua Christian Hospital, Changhua, Taiwan.
REVIEW RETURNED	18-Dec-2019

GENERAL COMMENTS	The authors have made a nice revision.

REVIEWER	Ana Timoteo Cardiology Department, Santa Marta Hospital, Lisbon, Portugal
REVIEW RETURNED	13-Dec-2019

GENERAL COMMENTS	There are some limitations in this paper, that were mentioned by all the reviewers. The authors tried to improve data analysis with our suggestions. Despite the remaining of some limitations, I think it is suitable for publications.

VERSION 2 – AUTHOR RESPONSE

Q1. Title: I would suggest to delete "The" as first word in the title.

Reply 1: Thanks for your comment. It has been revised.

Q2. Item 2. suggest to swap "should be" for "were"

Reply 2: Thanks for your comment. It has been revised.

Q3. Pg. 10 Line 12: "However, there is no significant difference were found"

Rephrase: However, no significant differences were found?

Reply 3: Thanks for your comment. It has been revised.

Q4. Pg 10 lines 22-24 : Notably, despite no significant effect of VV on the survival of patients with

hypertension, hyperlipidemia or coronary artery disease was observed?, patients with both VV and diabetes presented a 1.50 times higher risk of mortality compared with those without VV (adjusted HR: 1.50; 95% CI: 1.05–2.15; p = 0.0254).
Reply 4: Thanks for your comment. It has been revised.

Q5. Pg 11. Line 3: under the control of = adjusted for?
Reply 5: Thanks for your comment. It has been revised.

Q6. Pg 11, Line 12: In patients with VV, 3-, 6-, and 9-year MACE-free rates were 91.17% (91.2%), 84.99% (85%), and 79.27% (79.3%),
Reply 6: Thanks for your comment. It has been revised.

Q7. Pg 11, lines 18-19: including DVT and PE (Grade 3: adjusted HR: 38.42 (38.4); 95% CI, 16.38 (16.4)–90.13(90.1); p < 0.0001) (Table 4).
Pg 12, lines 10-11: The calculated kappa score between CEAP stages and grading severity is 0.918 (0.92) (95%CI = [0.878(0.88), 0.957(0.96)]).
Pg 12, line 22: treated VV patients were found 1.36 times risks of mortality (adjusted HR (95%CI) = 1.362(136) (1.18, 1.57), p-value<0.0001)
Reply 7: Thanks for your comment. It has been revised.

Q8. Page 14, lines 13-16: Please rephrase the following insert: " Similarly, Lohr et al also reviewed that Although a lower grade of VV (CEAP 2-3) has been observed in 50.5% of females and in 30.1% of males, a higher grade of VV with trophic skin changes (CEAP 4-6) were found in 2.8% of females and 5.4% of males.20 Also, DVT was more common in males compared with females (11.3% vs 7.8%).20"
Reply 8: Thanks for your comment. It has been rephrased.

Q9. The discussion could profit from editorial revision.
Tables 2, 3,4. And supplement tables: Please remove third decimal throughout in the table. (1.431 (1.247, 1.643) should be 1.43 (1.25, 1.64) etc.
Figure 1,2: There is a 12 year follow up , the KM plots show 15 years and 0 survivors. I would suggest limiting the KM plots to the observed follow up period of 12 years.
Reply 9: Thanks for your comment. It has been revised.

We will be extremely grateful if you may understand the limitation of this study while it remains showing important information regarding the possible under-evaluation of the risks of VVs.

VERSION 3 – REVIEW

REVIEWER	Dr. A.J. ten Cate-Hoek Maastricht University Medical Centre, the Netherlands.
REVIEW RETURNED	10-Mar-2020

GENERAL COMMENTS	All questions and remarks have been addressed to my satisfaction.